

ESTIMACIÓN DE MODELOS DE MORTALIDAD ESTOCÁSTICA PARA CHILE

ESTIMATION OF STOCHASTIC MORTALITY MODELS FOR CHILE

Pablo Andrés Moyano Silva¹, Ana María Pérez Marín^{2*} y Miguel Santolino²

¹*Depto. Matemática y Ciencia de la Computación, Universidad de Santiago
de Chile*

Santiago, Chile

²*Depto. Econometría, Estadística y E.A., Riskcenter-IREA, Universitat de
Barcelona*

Barcelona, España

**Autora para correspondencia*

Resumen

En este artículo se estiman los modelos de Lee-Carter, Renshaw-Haberman y Cairns-Blake-Dowd, expuestos en el marco de modelos de mortalidad estocástica edad-periodo-cohorte generalizados, con datos de mortalidad de Chile. El modelo de Lee-Carter con distribución binomial resulta ser el más adecuado para describir la evolución de la mortalidad en Chile. Se muestra una aplicación de la construcción de tablas de mortalidad para el cálculo de primas en seguros de vida, y se comparan con las obtenidas utilizando la tabla de mortalidad M95 (publicada por el regulador en Chile). Se concluye que las tablas M95 parecen adecuadas para garantizar la solvencia. Finalmente, se calcula la esperanza de vida al nacer en el año 2019 y en el año 2050 por sexo.

Palabras clave: riesgo longevidad, modelo mortalidad, regresión no-lineal, pensiones.

A. M. Pérez-Marín y M. Santolino agradecen la ayuda recibida del Ministerio de Ciencia e Innovación, PID2019-105986GB-C21.

Abstract

In this article we fit the Lee-Carter, Renshaw-Haberman and Cairns-Blake-Dowd models, presented in the framework of generalised age-period-cohort stochastic mortality models, to mortality data of Chile. The Lee-Carter model with binomial distribution is the most appropriate model to describe the evolution of mortality for the population of Chile. We show an application of the creation of mortality tables for premium calculation in life insurance, and we compare the results with those obtained by using the M95 mortality table (published by the regulator in Chile). The M95 tables seem to be appropriate to guarantee solvency. Finally, the life expectancy at birth in 2019 and 2050 are calculated by sex.

Keywords: longevity risk, mortality model, non-linear regression, pensions.

1. Introducción

En las últimas décadas se ha constatado un progresivo aumento en la esperanza de vida de la población de Chile, así como en el resto de países de su entorno. La edad media de fallecimiento de la población en Chile en el año 1990 era de 68,8 años y en 2010 se situaba en 78,4 años (CEPAL, 2017). Aproximadamente en el mismo periodo, la tasa global de fecundidad pasó de 2,36 hijos por mujer en edad fértil (año 1992) a 1,87 hijos (año 2010) (DEIS, 2020). La mayor longevidad unida a una menor tasa de fecundidad implica un envejecimiento progresivo de la sociedad chilena, lo que representa importantes desafíos para los sistemas públicos y privados de salud y previsión social. Además, Chile se sitúa en puestos elevados en el ranking de países de América Latina y el Caribe respecto a la eficiencia del gasto sanitario a la hora de mejorar la esperanza de vida y reducir la mortalidad infantil (Melgen-Bello y García-Prieto, 2017).

Todo indica que las mejoras de longevidad continuarán en el futuro más inmediato. En 2010 el 13% de la población tenía más de 60 años. En 2018 este porcentaje ascendió al 16,3% y en 2050 se espera que pueda llegar al 30,7% (Arnold, Herrera, Massad, y Thumala, 2018). A su vez, la relación de dependencia, que mide el cociente entre número de personas inactivas (de 0 a 14 años más la población de 60 y más) y las personas activas (CEPAL, 2000) va en aumento en Chile. Mientras en el año 1970 había 9,2 adultos mayores por cada cien habitantes en edad laboral, para la mitad del presente siglo este índice puede alcanzar 34,91 (Jiménez, 2008).

El objetivo del estudio es estimar la evolución futura de la mortalidad en la población chilena utilizando los principales modelos estocásticos de mortalidad existentes, y analizar si las tablas de mortalidad que actualmente se utilizan en los seguros de vida son adecuadas según las proyecciones de mortalidad realizadas por estos modelos. En este artículo se ajustan los principales modelos estocásticos de mortalidad para Chile para el periodo 1992-2008. Se comparan estos modelos desde el punto de vista de su bondad de ajuste y capacidad de predicción. Una vez seleccionados los modelos de mortalidad con mejor comportamiento, las proyecciones de mortalidad futura se utilizan para la construcción de tablas de mortalidad generacionales. Se muestra la aplicación de las tablas de mortalidad estimadas en el cálculo de primas en los seguros de vida. Mediante la definición de distintos escenarios, se compara la prima pura obtenida cuando se utilizan las tablas de mortalidad estimadas en este artículo con la prima pura obtenida cuando se aplican las tablas M95, que son las tablas vigentes en Chile desde el año 2000 y su uso es obligatorio por parte de las compañías de seguro para el cálculo de las reservas matemáticas en los seguros de vida. Esta comparación nos debe permitir analizar la suficiencia de las primas obtenidas mediante las tablas M95. Por último, se utilizan las predicciones de las tasas de mortalidad de la población de Chile para realizar estimaciones de la esperanza de vida al nacer.

En este artículo se considera la familia de modelos de mortalidad estocástica edad-periodo-cohorte generalizados definidos en Villegas, Millosovich y Kaishev (2017) para modelizar la evolución de la mortalidad. En particular, se trabaja con los modelos de Lee-Carter (Lee y Carter, 1992), Renshaw-Haberman (Renshaw y Haberman, 2003; Haberman y Renshaw, 2011) y Cairns-Blake-Dowd (Blake, Cairns y Dowd, 2006). Aunque este artículo se centra en los tres principales modelos de mortalidad utilizados en la literatura, otras propuestas interesantes de modelización se engloban dentro de los modelos edad-periodo-cohorte. Este es, por ejemplo, el caso del modelo de Plat (Plat, 2009) que combina el modelo CBD y el modelo Lee-Carter de forma que se capture el efecto cohorte. Para una recopilación de estos modelos, véase Villegas et al. (2017).

Este artículo se organiza de la siguiente forma. En la siguiente sección se presenta la metodología utilizada, describiendo el marco teórico de la familia de modelos de mortalidad estocástica edad-periodo-cohorte generalizados. En la sección 3 se presentan los datos y los resultados de la estimación de los modelos, comparándolos entre ellos. En la sección 4 se presenta una aplicación actuarial al cálculo de primas de seguros de vida en diferentes escenarios. También se muestran los resultados del cálculo de las predicciones de la esperanza de vida futura al nacer. Finalmente, en la sección 5 se

describen las principales conclusiones.

2. Metodología

Se considera que los individuos forman un grupo homogéneo, en el sentido de que el comportamiento estadístico de su edad de fallecimiento es idéntico. Esto permite estudiar el comportamiento probabilístico de un individuo genérico, y obtener conclusiones para todo el grupo. Cuando esta hipótesis no resulta aceptable, se divide a los individuos en grupos cuya homogeneidad no sea puesta en duda. En este caso, se clasifica a los individuos por sexo.

Se considera que las variables que describen las edades de fallecimiento de los individuos son estadísticamente independientes. Esto implica que las probabilidades para la edad de fallecimiento de un individuo no dependen de la edad de fallecimiento de otro individuo. Se utiliza la misma notación que Ayuso, Corrales, Guillén, Pérez-Marín y Rojo (2007). El subíndice x hace referencia a la edad, y el subíndice t al año calendario. De este modo, la variable aleatoria D_{xt} denota el número de muertes en una población de edad x en el año calendario t . El número observado de fallecimientos se representa como d_{xt} .

La probabilidad de fallecimiento a un horizonte temporal de un año, para un individuo de edad x en el año calendario t se expresa como q_{xt} . La fuerza de la mortalidad (o tanto instantáneo de fallecimiento) para un individuo de edad x en el año calendario t , es μ_{xt} y se define como

$$\mu_{xt} = \lim_{h \rightarrow 0} \frac{h q_{xt}}{h}$$



donde $h q_{xt}$ es la probabilidad de fallecimiento a un horizonte temporal h . Así, para un h pequeño, la probabilidad de fallecer de este individuo entre el periodo t y $t + h$ es aproximadamente $\mu_{xt} \cdot h$.

La función censal de supervivencia o exposición al riesgo E_{xt} recoge el promedio de individuos vivos a lo largo de la edad x en el año calendario t . E_{xt}^0 corresponde a la exposición inicial al riesgo, que es la cantidad de individuos de edad x al inicio del año calendario t . La exposición central al riesgo E_{xt}^c hace referencia a la población viva a la mitad del año calendario (Macías y Santolino, 2018). Si se asume mortalidad uniforme a lo largo del año, se cumple que

$$E_{xt}^0 \approx E_{xt}^c + \frac{1}{2}d_{xt}.$$



La probabilidad de fallecimiento a un horizonte temporal de un año puede estimarse como $q_{xt} = d_{xt}/E_{xt}^0$. La tasa central de mortalidad se denota por m_{xt} y es la razón entre el número de fallecimientos de edad x y la población viva de edad x a la mitad del año calendario t . Puede estimarse como $m_{xt} = d_{xt}/E_{xt}^c$. Se asume que la fuerza de mortalidad es constante para cada edad y año calendario. Es decir, μ_{xt} es constante de la edad x a la edad $x + 1$ y del año t al año $t + 1$. Bajo este supuesto, la fuerza de la mortalidad μ_{xt} coincide con la tasa central de mortalidad m_{xt} .

2.1. Principales modelos de mortalidad bajo el enfoque de los modelos GAPC

En Currie (2016) se muestra que los modelos de mortalidad más utilizados en la actualidad pueden expresarse utilizando la estructura de modelos (no)lineales generalizados, y en esta línea Villegas et al. (2017) define la familia de modelos de mortalidad estocástica edad-periodo-cohorte generalizados (GAPC por sus siglas en inglés). Formalmente, un modelo de mortalidad estocástica GAPC consta de cuatro componentes:

- Componente aleatorio: El número de muertes D_{xt} sigue una distribución Poisson con tasa $E_{xt}^c\mu_{xt}$

$$D_{xt} \sim \text{Poisson}(E_{xt}^c\mu_{xt})$$



o sigue una distribución Binomial con parámetros E_{xt}^0, q_{xt}

$$D_{xt} \sim \text{binomial}(E_{xt}^0, q_{xt})$$



donde $\mathbb{E}(D_{xt}/E_{xt}^c) = \mu_{xt}$ y $\mathbb{E}\left(\frac{D_{xt}}{E_{xt}^0}\right) = q_{xt}$, respectivamente.

- Componente sistemático: El predictor lineal η_{xt} dado por:

$$\eta_{xt} = \alpha_x + \sum_{i=1}^N \beta_x^{(i)} k_t^{(i)} + \beta_x^{(0)} \gamma_{t-x}$$



recoge los efectos de la edad x , del año calendario t y de la cohorte $c = t - x$. α_x es una función que depende de la edad y que captura la forma general de la mortalidad, $N \geq 0$ es un entero que indica el número de tramos de edad utilizados para describir la tendencia de la mortalidad, $k_t^{(i)}$, $i =$

$1, \dots, N$ son los efectos de periodo que contribuyen a especificar la tendencia de mortalidad y $\beta_x^{(i)}$ modula su efecto a través de las edades. El término γ_{t-x} explica el efecto de cohorte y $\beta_x^{(0)}$ modula su impacto en las diferentes edades.

- Componente de enlace: Una función de enlace g que relaciona la esperanza matemática de la variable dependiente con el predictor lineal, esto es

$$g\left(\mathbb{E}\left(\frac{D_{xt}}{E_{xt}^c}\right)\right) = \eta_{xt}. \quad \text{🗨️}$$

En el caso de utilizar las funciones de enlace canónicas, g es la función de enlace *log* si se considera la distribución de Poisson o *logit* si se trabaja con la distribución binomial.

- Conjunto de restricciones: La mayoría de los modelos de mortalidad estocástica necesitan restricciones en los parámetros para garantizar una estimación única de los mismos. Estas restricciones se aplican a través de una función v que toma un vector arbitrario de parámetros

$$\theta := (\alpha_x, \beta_x^{(1)}, \dots, \beta_x^{(N)}, k_t^{(1)}, \dots, k_t^{(N)}, \beta_x^{(0)}, \gamma_{t-x}) \quad \text{🗨️}$$

y lo transforma en el vector

$$v(\theta) = \tilde{\theta} = (\tilde{\alpha}_x, \tilde{\beta}_x^{(1)}, \dots, \tilde{\beta}_x^{(N)}, \tilde{k}_t^{(1)}, \dots, \tilde{k}_t^{(N)}, \tilde{\beta}_x^{(0)}, \tilde{\gamma}_{t-x})$$

que satisface las restricciones del modelo sin efecto sobre el predictor η_{xt} . 🗨️

2.1.1. Modelo Lee-Carter

Lee y Carter (1992) propusieron un sencillo modelo para describir la evolución de la mortalidad. El modelo Lee-Carter ha sido utilizado para modelizar y pronosticar la mortalidad en diversos países como, por ejemplo, en Estados Unidos (Lee y Carter, 1992), Chile (Lee y Rofman, 1994), Bélgica (Brouhns, Denuit y Vermunt, 2002) y España (Betzen, 2010; Debón, Martínez, Montes y Moshuk, 2015; Santolino, 2020), entre otros. Es un modelo bilineal en las variables edad (x) y año calendario (t) y se expresa de la siguiente manera:

$$\ln(m_{x,t}) = a_x + \beta_x^{(1)} \cdot k_t^{(1)} + \varepsilon_{x,t} \quad \text{💬}$$

donde $\varepsilon_{x,t}$ es el término de error con media 0 y varianza σ_ε^2 y representa todos aquellos aspectos no capturados por el modelo. En general, se asume que ε sigue una distribución gaussiana. Lee y Carter (1992) propusieron calibrar el modelo minimizando el error cuadrático medio mediante la descomposición en valores singulares.

El modelo Lee-Carter en el marco de los modelos GAPC fue implementado en Brouhns et al. (2002), donde se asume que el número de muertes D_{xt} tiene distribución Poisson, con función de enlace *log* y con media μ_{xt} multiplicado por la exposición al riesgo E_{xt}^c . Posteriormente, fue extendido al caso de distribución binomial. Así, el predictor viene dado por:

$$\eta_{xt} = \alpha_x + \beta_x^{(1)} k_t^{(1)}. \quad \text{💬}$$

Los parámetros de la ecuación anterior pueden ser transformados de la siguiente manera:

$$(\alpha_x, \beta_x^{(1)}, k_t^{(1)}) \rightarrow \left(\alpha_x + c_1 \beta_x^{(1)}, \frac{1}{c_2} \beta_x^{(1)}, c_2 (k_t^{(1)} - c_1) \right) \quad c_1, c_2 \in \mathbb{R}, c_1 \neq 0, c_2 \neq 0 \quad \text{💬}$$

dejando invariante el valor de η_{xt} . Para asegurar la identificación (e interpretabilidad) de los parámetros, se establecieron las siguientes restricciones a los parámetros

$$\sum_x \beta_x^{(1)} = 1 \quad \sum_t k_t^{(1)} = 0 \quad \text{💬}$$

que se obtienen mediante

$$c_1 = \frac{1}{n} \sum_t k_t^{(1)} \quad c_2 = \sum_x \beta_x^{(1)}.$$

2.1.2. Modelo Renshaw-Haberman

Renshaw y Haberman (2006) extienden el modelo original de Lee-Carter e incorporan un parámetro para capturar el efecto global de la cohorte $t - x$. El efecto de cohorte captura modificaciones en la mortalidad que una generación experimenta respecto a las generaciones previas y/o posteriores y que se prolonga a lo largo del tiempo en dicha generación. Estos cambios entre cohortes se deben a múltiples causas como pueden ser, por ejemplo, guerras, epidemias o mejoras en la mortalidad infantil.

En el marco de los modelos GAPC, el número de muertes D_{xt} se distribuye según una Poisson o binomial con predictor lineal

$$\eta_{xt} = \alpha_x + \beta_x^{(1)} k_t^{(1)} + \beta_x^{(0)} \gamma_{t-x}. \quad \text{💬}$$

Más adelante, los autores proponen fijar $\beta_x^{(0)} = 1$, pues de este modo se resuelven ciertos problemas de estabilidad del modelo general (Haberman y Renshaw, 2011). En este trabajo se sigue esta modelización.

El predictor queda invariante bajo la transformación de parámetros

$$(\alpha_x, \beta_x^{(1)}, k_t^{(1)}, \gamma_{t-x}) \rightarrow \left(\alpha_x + c_1 \beta_x^{(1)} + c_2, \frac{1}{c_3} \beta_x^{(1)}, c_3(k_t^{(1)} - c_1), \gamma_{t-x} - c_2 \right)$$

con $c_1, c_2, c_3 \in \mathbb{R}$, $c_1 \neq 0, c_2 \neq 0, c_3 \neq 0$. El conjunto de restricciones para los parámetros es

$$\sum_x \beta_x^{(1)} = 1 \quad \sum_t k_t^{(1)} = 0 \quad \sum_{c=t_1-x_k}^{t_n-x_1} \gamma_c = 0 \quad \text{💬}$$

que se obtiene imponiendo:

$$c_1 = \frac{1}{n} \sum_t k_t^{(1)} \quad , \quad c_2 = \frac{1}{n+k-1} \sum_{c=t_1-x_k}^{t_n-x_1} \gamma_c \quad , \quad c_3 = \sum_x \beta_x^{(1)}.$$

2.1.3. Modelo Cairns-Blake-Dowd (CBD)

Blake et al. (2006) proponen un predictor con dos tramos de edad ($N = 2$) para describir la tendencia de la mortalidad. Sus respectivos efectos son $k_t^{(i)}$, $i = 1, 2$, con parámetros $\beta_x^{(1)} = 1$ y $\beta_x^{(2)} = x - \bar{x}$, respectivamente, donde \bar{x} es el promedio de edad. En este caso, no existe función de la edad α_x ni efectos de cohorte γ_{t-x} . El predictor del modelo CBD es

$$\eta_{xt} = k_t^{(1)} + (x - \bar{x}) k_t^{(2)} \quad \text{💬}$$

Una característica que contrasta con los modelos anteriores, es que en este modelo la edad es una variable explicativa. No se dan problemas de identificación por lo que no se requiere de un conjunto de restricciones para los parámetros.

2.2. Calidad del ajuste

La bondad de ajuste de los modelos de mortalidad usualmente se analiza mediante la inspección de los residuos del modelo ajustado. Los patrones regulares en los residuos indican la incapacidad del modelo para describir todas las características de los datos de manera adecuada. Los modelos GAPC son modelos (no-)lineales generalizados cuyos parámetros se estiman por máxima verosimilitud. En el contexto de los modelos lineales generalizados, la devianza residual es de gran utilidad para analizar la bondad del ajuste del modelo. Estos residuos también pueden calcularse para los modelos GAPC.

Otras medidas de bondad de ajuste obtenidos a partir de la función de verosimilitud son el criterio de información de Akaike y el criterio de información Bayesiano. Cuando se evalúa la bondad de ajuste de diferentes modelos, es esperable que los modelos con más parámetros proporcionen un mejor ajuste a los datos. Los criterios de información de Akaike y Bayesiano asignan un valor a cada modelo en función de su verosimilitud y de la cantidad de parámetros que tienen, medida que genera un balance entre la bondad de ajuste del modelo y su simplicidad. Matemáticamente se definen como

$$AIC = 2\nu - 2\log(\mathcal{L})$$



$$BIC = \nu \log K - 2\log(\mathcal{L})$$

donde ν es el número de parámetros estimados del modelo, \mathcal{L} es el máximo valor de la función de verosimilitud del modelo y K es el número de observaciones. Se prefieren los modelos con un menor valor de AIC y BIC.

Por último, resulta interesante saber qué modelo estima mejor el número de muertes, y para este propósito se utiliza como medida de comparación entre los modelos estimados, el error cuadrático medio (ECM) y el error porcentual absoluto medio (EPAM), que miden la diferencia entre el número de muertes observadas d_{xt} y estimadas \hat{d}_{xt} .

El error cuadrático medio mide las diferencias en términos de distancia euclídea entre los valores observados y ajustados. Se define como:

$$ECM(\hat{d}_{xt}) = \sqrt{\sum_x \frac{(d_{xt} - \hat{d}_{xt})^2}{n}}, \quad t = t_1, \dots, t_n.$$



A su vez, el error porcentual absoluto medio (EPAM) mide la distancia relativa entre los valores observados y ajustados. Se define como

$$EPAM(\hat{d}_{xt}) = \frac{\sum_x \frac{|d_{xt} - \bar{d}_{xt}|}{d_{xt}}}{n}, \quad t = t_1, \dots, t_n.$$



Cabe señalar que tanto el ECM como el EPAM pueden ser utilizados aunque el modelo no se haya estimado por máxima verosimilitud, como en este trabajo. Este es el caso del modelo de Lee-Carter gaussiano estimado por mínimo cuadrados, de acuerdo a su versión original.

2.3. Proyecciones con modelos de mortalidad

Para realizar predicciones, se debe modelizar el comportamiento de los parámetros $k_t^{(i)}$, $i = 1, \dots, N$ y γ_{t-x} , pues proyectando los valores futuros de estos parámetros se pueden obtener proyecciones de las tasas de mortalidad. Para modelar los efectos de periodo $k_t^{(i)}$, $i = 1, \dots, N$, cuando $i > 1$, se puede considerar que los efectos de periodo siguen un paseo aleatorio multivariante con constante. Una segunda alternativa, es asumir que cada uno de los efectos de periodo sigue un modelo ARIMA univariante, es decir, se asume que el efecto de periodo i -ésimo $k_t^{(i)}$ sigue un $ARIMA(p_i, d_i, q_i)$ con constante,

$$\Delta^{d_i} k_t^{(i)} = \delta_0^{(i)} + \phi_1^{(i)} \Delta^{d_i} k_{t-1}^{(i)} + \dots + \phi_{p_i}^{(i)} \Delta^{d_i} k_{t-p_i}^{(i)} + \xi_t^{(i)} + \delta_1^{(i)} \xi_{t-1}^{(i)} + \dots + \delta_{q_i}^{(i)} \xi_{t-q_i}^{(i)}$$



donde $\Delta^{d_i} k_t^{(i)} = k_t^{(i)} - k_{t-d}^{(i)}$ es el operador de diferencias, $\delta_0^{(i)}$ es la constante, $\phi_1^{(i)}, \dots, \phi_{p_i}^{(i)}$ son los coeficientes de la parte autorregresiva con $\phi_{p_i}^{(i)} \neq 0$, $\delta_1^{(i)}, \dots, \delta_{q_i}^{(i)}$ son los coeficientes de la parte media móvil con $\delta_{q_i}^{(i)} \neq 0$ y $\xi_t^{(i)}$ es un ruido blanco gaussiano con varianza $\sigma_\xi^{(i)}$.

Por su parte, para modelar el efecto de cohorte γ_{t-x} , se asume que éste sigue un proceso $ARIMA(p, d, q)$ con constante, que es independiente de los efectos de periodo (Renshaw y Haberman, 2006):

$$\Delta^d \gamma_{t-x} = \delta_0 + \phi_1 \Delta^d \gamma_{t-x-1} + \dots + \phi_p \Delta^d \gamma_{t-x-p} + \epsilon_{t-x} + \delta_1 \epsilon_{t-x-1} + \dots + \delta_q \epsilon_{t-x-q}$$



donde ϵ_{t-x} es un ruido blanco gaussiano con varianza σ_ϵ .

Se utiliza la metodología de Box-Jenkins de series temporales para modelar el comportamiento de los parámetros $k_t^{(i)}$, $i = 1, \dots, N$ y γ_{t-x} . En términos generales, la metodología de Box-Jenkins consiste en la identificación, estimación y comprobación, para encontrar un proceso ARIMA que represente el mejor ajuste a una serie temporal de valores.

3. Estimación de los modelos de mortalidad para Chile

Los datos de mortalidad de Chile que se utilizan en este trabajo han sido obtenidos de la base de datos The Human Mortality Database (www.mortality.org) y abarcan edades entre 0 y 110 años, en el periodo de 1992 a 2008, diferenciando por sexo. Todos los datos fueron tratados mediante el software R (R Core Team, 2020).

La figura 1 muestra el logaritmo de la tasa central de mortalidad por edad, para hombres y mujeres, entre los años 1992 y 2008. Aunque el periodo de tiempo es breve, se puede observar en ambos gráficos que la mortalidad disminuye con el paso de los años.

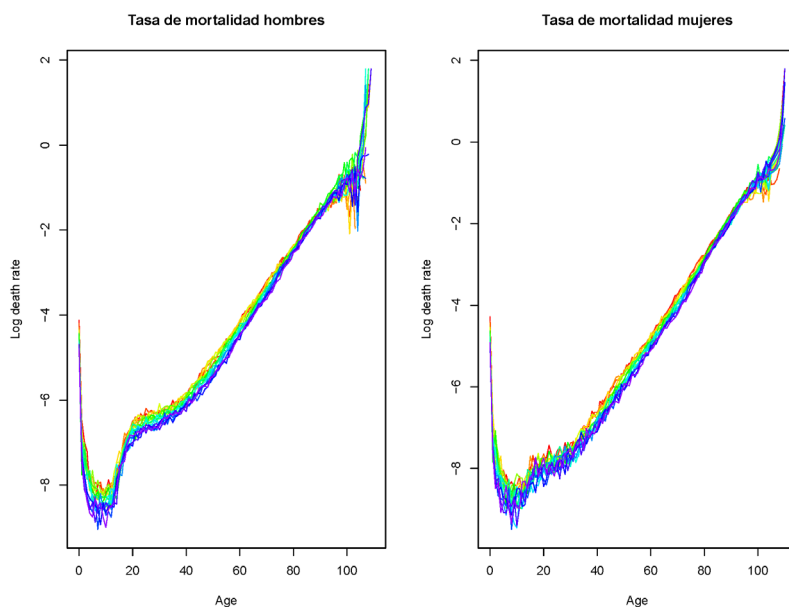


Figura 1. Tasa central de mortalidad para hombres y mujeres de Chile entre los años 1992 y 2008.

El comportamiento de la mortalidad es parecido en ambos gráficos. En términos generales, se observa una mortalidad alta en los recién nacidos, que luego cae rápidamente hasta un mínimo en torno a los 10 años, para luego volver a crecer paulatinamente. Podemos también observar diferencias entre hombres y mujeres, debido a un aumento de la mortalidad en el caso de los hombres entre los 20 y los 30 años. Esta diferencia en la mortalidad podría estar causada por los fallecimientos por accidentes, no por causas naturales.

Se ajustan los modelos para edades entre 0 y 90 años, ya que el objetivo es aplicar las predicciones a la valoración de productos de seguros de vida. Se asume que a edades avanzadas las entidades aseguradoras no ofrecen este tipo de productos. En el último apartado en el que se calculan las esperanzas de vida, se amplía el ajuste hasta los 100 años. Los modelos GAPC fueron estimados mediante el paquete StMoMo (Villegas et al., 2017). Para ajustar el modelo original de Lee-Carter con error gaussiano se utiliza el paquete Demography.

3.1 Estimación del modelo Lee-Carter

En la figura 2 se muestran los coeficientes estimados para el modelo Lee-Carter gaussiano original. Se puede apreciar que el parámetro α_x captura la forma general de la mortalidad para cada edad. Se observa una mortalidad alta al inicio de la vida, que luego cae hasta un mínimo en torno a los 15 años, para luego volver a crecer. También se aprecia que la mortalidad de los hombres es mayor que la mortalidad de las mujeres. Esta distancia es mayor entre los 20 y los 30 años, posiblemente debido a los fallecimientos en accidentes de tránsito (joroba de los accidentes), tal y como se apuntó antes.

Por su parte, el parámetro β_x es siempre positivo tanto para hombres como para mujeres, y tiene un comportamiento irregular en ambos casos. Ello probablemente se deba a que intenta capturar las desviaciones del cambio de mortalidad para cada edad y también los efectos de cohorte (Macías y Santolino, 2018). La tendencia de la mortalidad es capturada por k_t , que para ambos sexos decrece, lo que indica que la mortalidad aumenta con la edad y disminuye con el tiempo calendario.

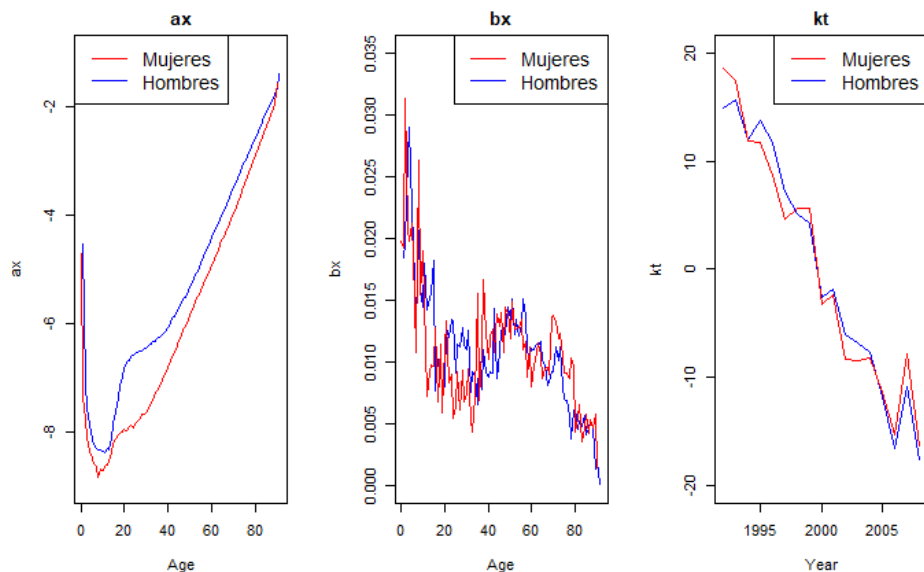


Figura 2. Estimación de los parámetros del modelo Lee-Carter Gaussiano

En el caso del modelo Lee-Carter en el entorno GAPC, se realizan dos ajustes, uno suponiendo que el número de muertes sigue una distribución binomial y otro con una distribución Poisson. Los coeficientes estimados son muy similares a los mostrados en la figura 2. En la tabla 1 se muestra el AIC y el BIC de los modelos Lee-Carter en el entorno GAPC. Se observa que los valores de bondad de ajuste AIC y BIC presentados en la tabla 1 son menores para el ajuste con distribución Binomial, en ambos sexos. Por ello, se puede concluir que para los datos utilizados, el modelo Lee-Carter que asume que el número de muertes sigue una distribución binomial presenta mejor ajuste que el mismo modelo que asume una distribución Poisson, tanto para hombres como para mujeres.

Tabla 1. Valores AIC y BIC del modelo Lee-Carter asumiendo que el número de muertes sigue una distribución Binomial, y asumiendo que el número de muertes sigue una distribución Poisson, ajustado para datos de mortalidad de hombres y mujeres de Chile.

Modelo	AIC	BIC
Lee-Carter Binomial Hombres	14538,34	15591,12
Lee-Carter Poisson Hombres	14588,89	15641,67
Lee-Carter Binomial Mujeres	13411,28	14464,07
Lee-Carter Poisson Mujeres	13450,09	14502,87

En las figuras 3 y 4 se representa la devianza residual según edad, año calendario y año de nacimiento del modelo de Lee-Carter con entorno binomial, por sexo. Se puede apreciar que para ambos sexos los residuos se encuentran mayoritariamente en el intervalo $[-2,2]$. Además, no se aprecian patrones de comportamiento en los gráficos. Estas características se interpretan como un ajuste adecuado del modelo.

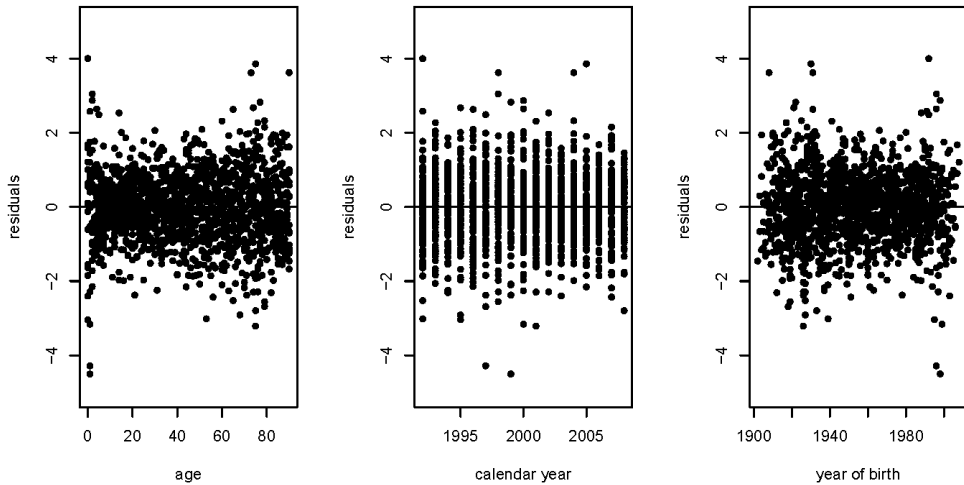


Figura 3. Residuos estandarizados basados en la devianza según edad, año calendario y año de nacimiento del modelo Lee-Carter con distribución binomial, para hombres.

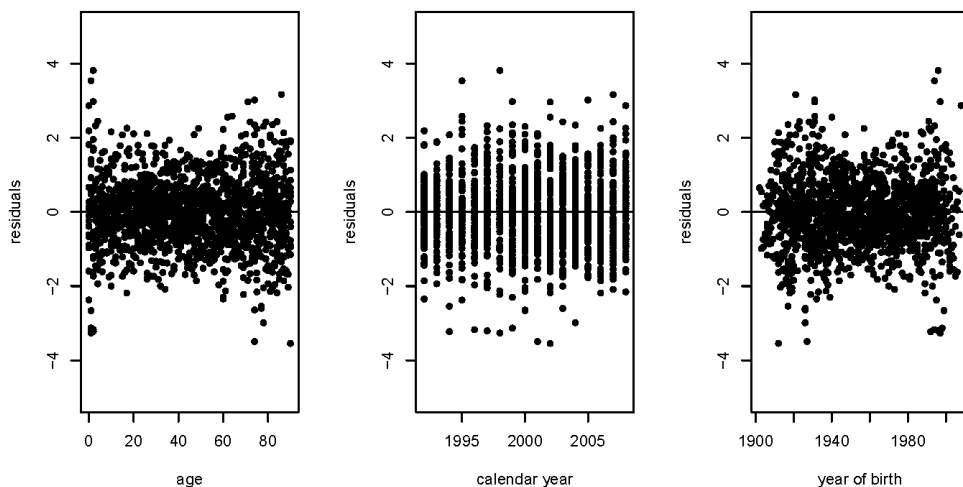


Figura 4. Residuos estandarizados basados en la devianza según edad, año calendario y año de nacimiento del modelo Lee-Carter con distribución binomial, para mujeres.

3.2. Estimación modelo de Renshaw-Haberman

A continuación, se muestran los valores AIC y BIC presentados en la tabla 2 para el modelo de Renshaw-Haberman, bajo los supuestos de que el número de muertes sigue una distribución Binomial y, por otra parte, Poisson.

Tabla 2. Valores AIC y BIC del modelo Renshaw-Haberman asumiendo que el número de muertes sigue una distribución binomial, y asumiendo que el número de muertes sigue una distribución Poisson, ajustado para datos de mortalidad de hombres y mujeres de Chile.

Modelo	AIC	BIC
Renshaw-Haberman Binomial Hombres	14109,29	15723,2
Renshaw-Haberman Poisson Hombres	14161,85	15775,76
Renshaw-Haberman Binomial Mujeres	13150,43	14764,34
Renshaw-Haberman Poisson Mujeres	13190,6	14804,51

Considerando que los valores AIC y BIC presentados en la tabla 2 son menores cuando se asume distribución binomial (tanto para hombres como para mujeres), se concluye que, para los datos utilizados, el modelo Renshaw-Haberman que asume distribución binomial proporciona un mejor ajuste.

La figura 5 presenta los gráficos de los parámetros estimados. Se aprecia que el parámetro α_x , que captura la forma general de la mortalidad para cada edad, y el efecto de periodo $k_t^{(1)}$, que captura la tendencia de la mortalidad, son similares a los obtenidos con el modelo de Lee-Carter. Se aprecia que la mortalidad aumenta con la edad y disminuye con los años calendarios. Por su parte, el parámetro β_x tiene un comportamiento distinto a lo estimado en el modelo de Lee-Carter y el efecto de cohorte γ_{t-x} tiene un comportamiento oscilante alrededor del cero. Aunque estos parámetros se han podido estimar, en este caso su interpretación es difusa.

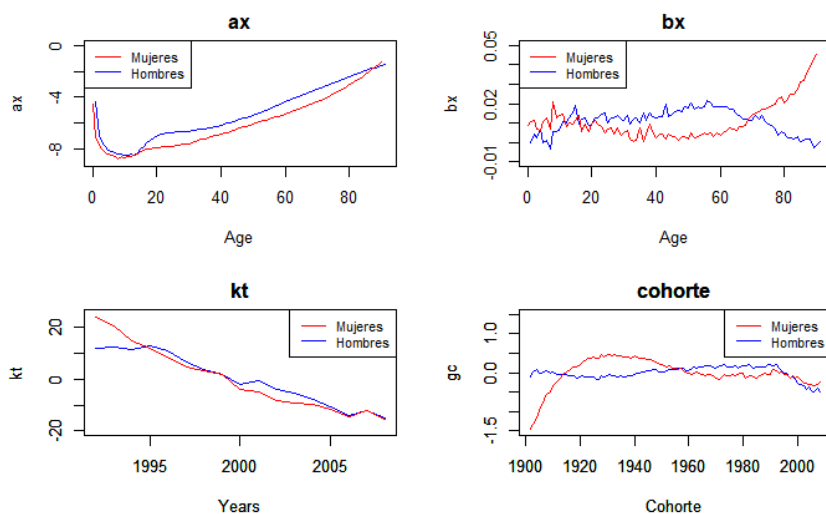


Figura 5. Estimación de los parámetros del modelo Renshaw-Haberman con distribución Binomial.

Las figuras 6 y 7 representan los residuos estandarizados basados en la devianza según edad, año calendario y año de nacimiento del modelo de Renshaw-Haberman con distribución binomial, por sexo. Se aprecia que para ambos sexos los residuos se concentran en el intervalo $[-2,2]$, no apreciándose patrones de comportamiento. El modelo presenta un adecuado

ajuste a los datos.

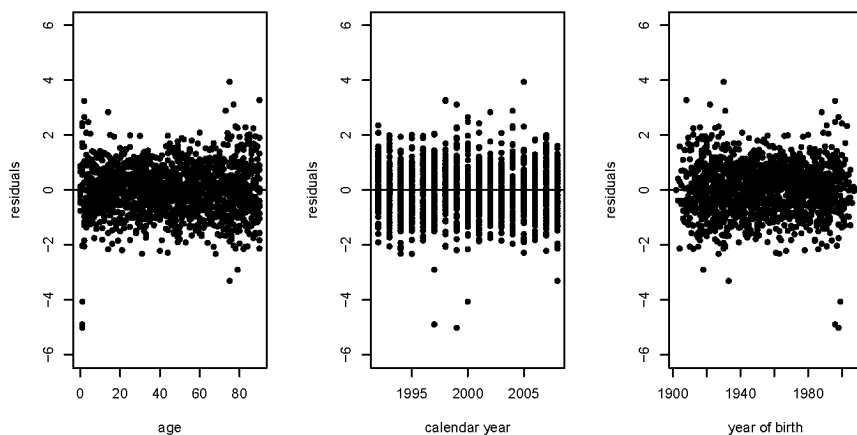


Figura 6. Residuos estandarizados basados en la devianza según edad, año calendario y año de nacimiento del modelo Renshaw-Haberman con distribución binomial, para hombres.

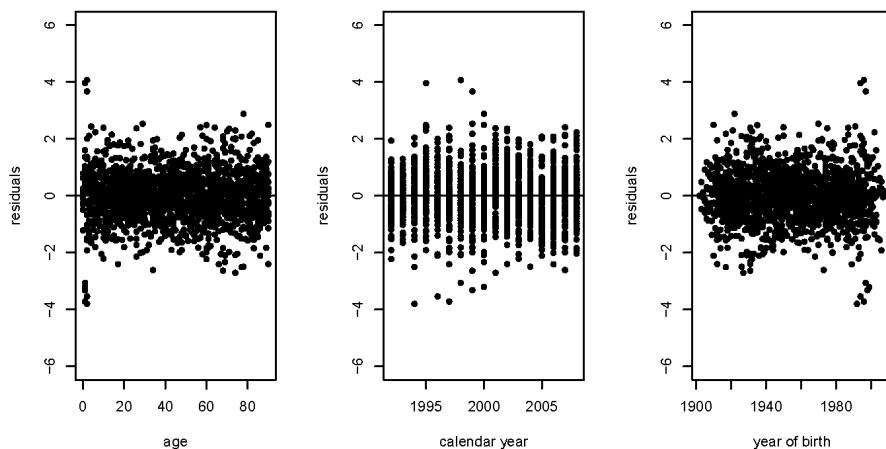


Figura 7. Residuos estandarizados basados en la devianza según edad, año calendario y año de nacimiento del modelo Renshaw-Haberman con distribución binomial, para mujeres.

En base al análisis de los residuos, se aprecia que el modelo de Renshaw-Haberman (asumiendo que las muertes se distribuyen como una

Binomial) presenta un adecuado ajuste para ambos sexos.

3.3. Estimación modelo de Cairns-Blake-Dowd

En la tabla 3 se muestran los valores de los criterios de información AIC y BIC cuando se ajusta el modelo de CBD. En el caso de la Poisson, se observa un mejor comportamiento, pero aún así son bastante más elevados que los obtenidos para los modelos de Lee-Carter y de Renshaw-Haberman.

Tabla 3. Valores AIC y BIC del modelo CBD asumiendo que el número de muertes sigue una distribución Binomial, y asumiendo que el número de muertes sigue una distribución Poisson, ajustado a datos de mortalidad de hombres y mujeres de Chile.

Modelo	AIC	BIC
Cairns-Blake-Dowd Binomial Hombres	207061,8	207243,5
Cairns-Blake-Dowd Poisson Hombres	203667,6	203849,3
Cairns-Blake-Dowd Binomial Mujeres	206962,8	207144,5
Cairns-Blake-Dowd Poisson Mujeres	204269,7	204451,4

En la figura 8 se aprecian los gráficos de los coeficientes estimados con el modelo CBD asumiendo una distribución Poisson. Se puede apreciar que, tanto para hombres como para mujeres, $k_t^{(1)}$ tiene una forma similar a los efectos de periodo de los modelos Lee-Carter. Por su parte los parámetros $k_t^{(2)}$ son crecientes para ambos sexos, lo cual indica que a menor edad las mejoras en la mortalidad son mayores que las mejoras en edades avanzadas.

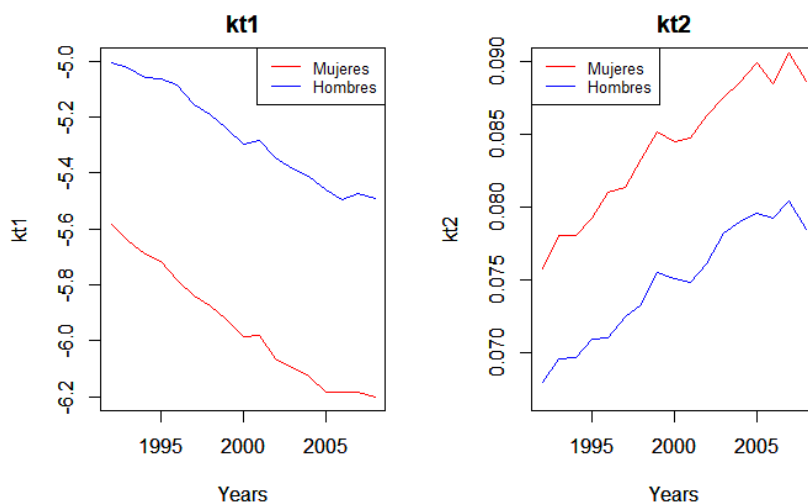


Figura 8. Estimación de los parámetros del modelo CBD con distribución Poisson.

Las figuras 9 y 10 muestran los residuos estandarizados basados en la devianza del modelo CBD con distribución Poisson, por sexo. Se aprecia, para ambos sexos, claros patrones de comportamiento, por lo que se concluye que el modelo no presenta un adecuado ajuste a los datos.

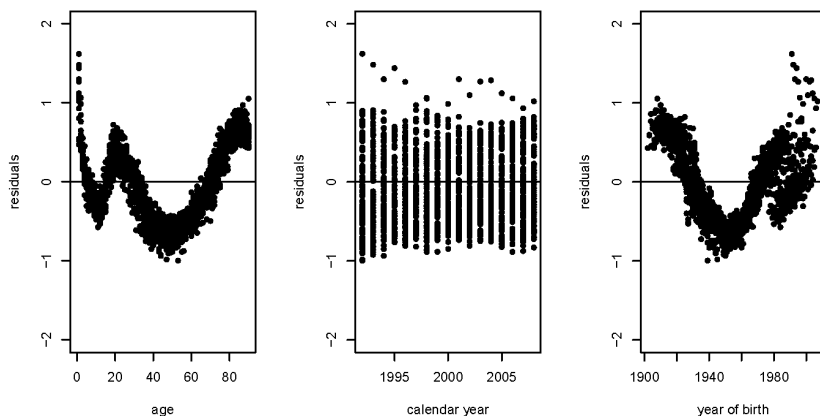


Figura 9. Residuos estandarizados basados en la devianza según edad, año calendario y año de nacimiento del modelo CBD con distribución Poisson, para hombres.

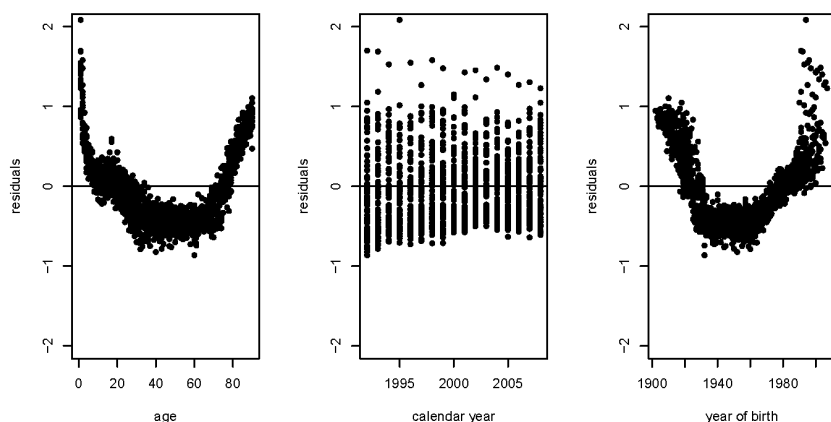


Figura 10. Residuos estandarizados basados en la devianza según edad, año calendario y año de nacimiento del modelo CBD con distribución Poisson, para mujeres.

3.4. Comparación entre modelos seleccionados

En la tabla 4 se presentan los valores del error cuadrático medio y del error porcentual absoluto medio para los modelos Lee-Carter y Renshaw-Haberman asumiendo distribución binomial para el número de muertes, del modelo CBD con distribución Poisson y del modelo Lee Carter original.

Tabla 4. Medidas de error para los modelos estimados con datos de mortalidad de Chile.

Modelo	ECM	EPAM
Lee-Carter Binomial Hombres	30,458	0,062
Renshaw-Haberman Binomial Hombres	24,4606	0,0568
Cairns-Blake-Dowd Poisson Hombres	190,5637	0,9465
Lee-Carter Gaussiano Hombres	33,4981	0,0651
Lee-Carter Binomial Mujeres	25,2083	0,08
Renshaw-Haberman Binomial Mujeres	20,0052	0,0744
Cairns-Blake-Dowd Poisson Mujeres	158,3856	2,1594
Lee-Carter Gaussiano Mujeres	39,0688	0,0847

Como era esperable, para ambos sexos el modelo CBD es el que peor resultados exhibe, por lo que deja de ser considerado en este estudio. A su vez, tanto para hombres como para mujeres, el modelo de Renshaw-Haberman es el que presenta menor error entre el número de muertes estimadas y su valor real. También se aprecia que existe una leve diferencia entre los desempeños de los modelos de Lee Carter con distribución binomial versus su versión original, siendo mejor el primero.

3.5. Desempeño de los modelos según sus pronósticos

Una vez analizada la capacidad de estimación de los modelos, se evalúa su capacidad de predicción, es decir, su capacidad para pronosticar la mortalidad futura (para años no utilizados en la estimación de los modelos). A solicitud de la Superintendencia de Valores y Seguros (actualmente Comisión para el Mercado Financiero), el Instituto Nacional de Estadísticas de Chile (INE) elaboró tablas de mortalidad para el periodo 2009 a 2012¹. Los modelos aquí considerados han sido ajustados en base a la experiencia de mortalidad del periodo que va de 1992 a 2008.

En esta sección se mide el desempeño de los modelos según los pronósticos generados comparándolos con las tablas elaboradas por el INE. Para realizar pronósticos con los modelos estimados, es necesario proyectar los parámetros temporales. Se utiliza la metodología de Box-Jenkins de series temporales para modelar el comportamiento de los parámetros temporales estimados. En el caso del modelo Lee-Carter gaussiano, tanto para hombres como para mujeres, el mejor proceso es un ARIMA (1,1,0) con constante. En el modelo Lee-Carter con distribución binomial, el mejor proceso para los hombres es un ARIMA (1,1,0) con constante, y, para las mujeres, un ARIMA(2,1,0) con constante. Por último, en el caso del modelo Renshaw-Haberman binomial, tanto para hombres como para mujeres, el mejor proceso para el efecto de periodo es un ARIMA (0,1,0) con constante, mientras que para el efecto de cohorte fue un ARIMA(2,2,2) para los hombres y un ARIMA(0,2,2) para las mujeres. En la tabla 5 se presentan los valores del error cuadrático medio y del error porcentual absoluto medio entre las tasas de mortalidad observadas por INE para los años 2009 a 2012 y las tasas de mortalidad pronosticadas en el mismo periodo de tiempo, para los modelos ajustados.

¹ Los datos están disponibles en web <http://www.svs.cl/portal/principal/605/w3-channel.html>

Tabla 5. Medidas de error para los pronósticos de los modelos estimados con datos de mortalidad de Chile

Modelo	ECM	EPAM
Lee-Carter Binomial Hombres	0.0031	0.0775
Renshaw-Haberman Binomial Hombres	0.003	0.0926
Lee-Carter Gaussiano Hombres	0.0026	0.0812
Lee-Carter Binomial Mujeres	0.0023	0.0804
Renshaw-Haberman Binomial Mujeres	0.0042	0.0932
Lee-Carter Gaussiano Mujeres	0.0025	0.0847

En base a ambas medidas, se observa que el modelo de Lee Carter con distribución binomial genera mejores pronósticos para los datos de mortalidad de las mujeres de Chile. Para el caso de los hombres, si se considera como medida el ECM, el modelo de Lee-Carter gaussiano es el que genera mejores predicciones. Si, en cambio, se considera como medida el EPAM, es el modelo Lee-Carter binomial. Ahora bien, en cada caso, los otros modelos tienen desempeños levemente distintos, por lo que se continuará trabajando con ellos.

4. Aplicación en seguros de vida y cálculo de la esperanza de vida al nacer

En esta sección se construyen tablas de mortalidad con los modelos ajustados y se utilizan para calcular la prima pura de seguros de vida, realizando la comparación de los resultados obtenidos. Para realizar dicha comparación, se definen dos escenarios hipotéticos, para los cuales se calculan las primas. Estos escenarios hacen referencia a personas que desean contratar un seguro de vida, por una cierta cantidad de años, para garantizar un capital a sus beneficiarios en caso de ocurrir su fallecimiento. Así, se define:

- Escenario 1: un hombre de 65 años en el año 2009, que desea contratar un seguro pagadero al final del año de fallecimiento de cuantía 60.000 euros, si su muerte ocurre en los próximos 15 años. El tipo de interés considerado es del 2% efectivo anual.
- Escenario 2: una mujer de 65 años en el año 2009, que desea contratar un seguro pagadero al final del año de fallecimiento de cuantía 60.000 euros, si su muerte ocurre en los próximos 15 años. El tipo de interés considerado es del 2% efectivo anual.

Para el cálculo de la prima en el ámbito de los seguros de vida, en Chile se utilizan las llamadas tablas M95. Estas tablas de mortalidad están vigentes

desde el año 2000 y fueron elaboradas por la Superintendencia de Valores y Seguros (actualmente esta institución se llama Comisión para el Mercado Financiero). Son de uso obligatorio para el cálculo de las reservas matemáticas de los seguros de vida. La tabla de mortalidad M95 H, es la utilizada en la constitución de reservas de pólizas de seguros de vida en que el asegurado sea de sexo masculino. Por otra parte, la tabla M95 M es la utilizada en la constitución de reservas de pólizas de seguros de vida en que el asegurado sea de sexo femenino.²

4.1. Tablas de mortalidad

En la tabla 6 se presentan los valores de q_{xt} para un hombre que en el año 2009 tiene 65 años. En las columnas 2, 3 y 4 se muestran las tablas de mortalidad ajustadas por el modelo de Lee-Carter con distribución binomial, considerando la estimación puntual en la columna 2, utilizando los valores inferiores del intervalo de confianza al 95% del pronóstico del efecto de periodo k_t en la columna 3 (que se denomina como Lee-Carter Binomial Low y se simboliza como LCBL) y utilizando los valores superiores del intervalo de confianza al 95% del pronóstico del efecto de periodo k_t en la columna 4 (que se denomina Lee-Carter Binomial Up y se simboliza como LCBU). Así mismo, en las columnas 5, 6 y 7 figuran las tablas de mortalidad ajustada por el modelo de Renshaw-Haberman con distribución binomial (RHB, RHBL y RHBV), y en las columnas 8, 9, 10 se recogen las tablas de mortalidad ajustada por el modelo de Lee-Carter gaussiano (LCG, LCGL y LCGU). Finalmente, en la columna 11 se recoge la tabla de mortalidad M95 H. Se puede apreciar que las menores tasas de mortalidad las presenta el modelo de Lee-Carter gaussiano y las tasas más altas el modelo de Renshaw-Haberman. Además los pronósticos del modelo de Renshaw-Haberman son los que más se aproximan a los valores de la tabla M95-H. También, se observa que las mayores diferencias entre los distintos modelos se dan en las edades más avanzadas.

² Disponible en

http://www.cmfchile.cl/institucional/mercados/ver_archivo.php?archivo=/web/compendio/cir/cir_1476_2000.pdf

Tabla 6. Escenario 1

Edad	LCB	LCBL	LCBU	RHB	RHBL	HBU	LCG	LCGL	LCGU	M95H
65	0,017	0,017	0,018	0,018	0,017	0,019	0,018	0,017	0,018	0,020
66	0,019	0,019	0,020	0,020	0,019	0,022	0,019	0,018	0,020	0,021
67	0,021	0,020	0,023	0,022	0,021	0,025	0,021	0,020	0,022	0,024
68	0,022	0,021	0,024	0,024	0,022	0,027	0,022	0,021	0,024	0,026
69	0,024	0,023	0,026	0,026	0,024	0,029	0,024	0,022	0,026	0,028
70	0,024	0,022	0,027	0,026	0,024	0,031	0,024	0,022	0,027	0,031
71	0,027	0,025	0,030	0,030	0,027	0,034	0,027	0,025	0,030	0,034
72	0,028	0,025	0,032	0,030	0,027	0,036	0,028	0,025	0,032	0,038
73	0,033	0,031	0,037	0,036	0,033	0,042	0,034	0,031	0,036	0,042
74	0,037	0,036	0,042	0,041	0,038	0,048	0,038	0,036	0,041	0,046
75	0,041	0,039	0,046	0,044	0,041	0,052	0,042	0,038	0,045	0,050
76	0,046	0,045	0,052	0,052	0,050	0,060	0,047	0,044	0,051	0,055
77	0,052	0,053	0,058	0,060	0,060	0,067	0,054	0,051	0,056	0,060
78	0,052	0,051	0,060	0,063	0,062	0,072	0,054	0,050	0,058	0,066
79	0,055	0,053	0,064	0,068	0,068	0,079	0,058	0,053	0,063	0,072

En la tabla 7 se presenta la misma información que en la tabla anterior, salvo que esta vez los datos son para una mujer que en el año 2009 tiene 65 años. A diferencia del escenario 1, aquí se puede apreciar que las menores tasas de mortalidad las presenta el modelo de Renshaw-Haberman y las tasas más altas el modelo de Lee-Carter gaussiano, aunque todos los modelos tienen tasas menores a los valores de la tabla M95-M. También se observa que las mayores diferencias entre los distintos modelos se dan en las edades más avanzadas, y para el caso de las mujeres esta diferencia es mayor que en el caso de los hombres.

Tabla 7. Escenario 2

Edad	LCB	LCBL	LCBU	RHB	RHBL	RHBU	LCG	LCGL	LCGU	M95M
65	0,010	0,010	0,010	0,010	0,010	0,010	0,010	0,010	0,011	0,011
66	0,011	0,010	0,011	0,010	0,010	0,011	0,011	0,010	0,012	0,012
67	0,011	0,011	0,012	0,011	0,010	0,012	0,012	0,011	0,013	0,013
68	0,012	0,012	0,013	0,012	0,011	0,013	0,013	0,012	0,014	0,015
69	0,012	0,012	0,013	0,011	0,010	0,013	0,012	0,011	0,014	0,016
70	0,013	0,012	0,014	0,012	0,010	0,013	0,013	0,011	0,015	0,018
71	0,014	0,013	0,015	0,012	0,010	0,014	0,014	0,012	0,017	0,020
72	0,016	0,015	0,017	0,013	0,011	0,016	0,016	0,014	0,019	0,022
73	0,017	0,016	0,018	0,013	0,011	0,016	0,017	0,015	0,020	0,024
74	0,019	0,018	0,021	0,014	0,012	0,018	0,020	0,017	0,023	0,027
75	0,022	0,022	0,024	0,016	0,012	0,020	0,023	0,020	0,027	0,030
76	0,024	0,024	0,027	0,016	0,013	0,021	0,025	0,022	0,029	0,033
77	0,027	0,026	0,030	0,016	0,013	0,022	0,028	0,024	0,032	0,036
78	0,028	0,027	0,031	0,016	0,012	0,023	0,029	0,024	0,034	0,040
79	0,030	0,029	0,034	0,017	0,012	0,024	0,032	0,027	0,038	0,045

4.2. Cálculo de primas puras

Para calcular las primas puras de los dos escenarios propuestos, se utilizarán las tablas de mortalidad presentadas en la sección anterior y se compararán los valores obtenidos con cada modelo con las primas obtenidas utilizando las tablas M95. Se denomina prima pura al valor esperado del coste real del riesgo que el asegurador asume.

El cálculo del valor actual actuarial en un seguro de vida pagadero al final del año de fallecimiento (VAA_s) se realiza mediante la formula

$$VAA_s = C \cdot A_{1_{x:\bar{d}|}} = \sum_{j=0}^{d-1} C \cdot (1+I)^{-(1+j)} \cdot {}_j p_x \cdot q_{x+j}$$



donde C es la cuantía a pagar en caso de fallecimiento, d es la temporalidad expresada en años, $A_{1_{x:\bar{d}|}}$ es el valor actual actuarial de una unidad monetaria pagadera al final del año de fallecimiento y con temporalidad d , I es el tipo de interés efectivo anual, ${}_j p_x$ es la probabilidad de un individuo de edad x que llegue vivo a la edad $x+j$, y q_{x+j} es la probabilidad de fallecimiento a un horizonte temporal de un año, para un individuo de edad $x+j$ (Macías y

Santolino, 2018).

En los dos escenarios antes definidos, la cuantía es de 60.000 euros, la temporalidad es de 15 años y el tipo de interés considerado es del 2% efectivo anual. En la tabla 8 se presenta el valor actual actuarial de los seguros de vida para ambos escenarios, utilizando las diferentes tablas de mortalidad que se han considerado. Se puede apreciar que, para el escenario 1, el valor actual actuarial del modelo de Renshaw-Haberman Up es el más alto, y el obtenido por el modelo de Lee-Carter Gaussiano Low es el más bajo. Además, destaca que solo el modelo de Renshaw-Haberman Up está por encima del valor actual actuarial calculado utilizando la tabla M95-H. Por su parte, se puede apreciar que, para el escenario 2, el valor actual actuarial utilizando cualquiera de las tablas de mortalidad consideradas es siempre inferior al correspondiente valor obtenido para el escenario 1. Además, es menor que el valor actuarial que se obtiene cuando se utiliza la tabla M95-M. Para este escenario 2, el valor actual actuarial del modelo Lee-Carter Gaussiano UP es el más próximo a los valores alcanzados por la tabla M95-M, mientras que el modelo de Renshaw-Haberman Low es el de menor valor actual actuarial.

Tabla 8. Valor actual actuarial

Modelos	Escenario 1	Escenario 2
Lee Carter Binomial	20052,82	11888,29
Lee-Carter Binomial-Low	19445,58	11478,20
Lee-Carter Binomial-Up	21925,92	12752,54
Renshaw-Haberman	21887,49	9257,81
Renshaw-Haberman-Low	21027,85	7958,71
Renshaw-Haberman-Up	24330,25	11106,41
Lee-Carter Gaussiano	20412,69	12252,60
Lee-Carter Gaussiano-Low	19224,97	10859,55
Lee-Carter Gaussiano-Up	21645,48	13806,27
M95	23476,02	15354,13

4.3. Estimación de la esperanza de vida al nacer

La esperanza de vida al nacer es un indicador sintético que resume la mortalidad a lo largo del tiempo de una cohorte hipotética de personas. Para su cálculo, utilizamos las proyecciones de probabilidad de que una persona nacida en un año concreto, sobreviva durante los cien años posteriores (u otro infinito actuarial) a su año de nacimiento. Ayuso, Bravo y Holzmann (2020) la

denomina la esperanza de vida de la cohorte, para diferenciarla de la esperanza de vida del periodo, que sería la edad media de fallecimiento para un año concreto.

En concreto, para su cálculo se utiliza la fórmula $e_{0,t} = \sum_{n=1}^{100} {}_n p_{0,t}$, donde ${}_n p_{0,t}$ es la probabilidad de que un individuo de cero años de edad en el año calendario t , sobreviva n años. Cuando n es mayor a uno, la probabilidad temporal de supervivencia se obtiene como el producto de probabilidades anuales, ${}_n p_{0,t} = p_{0,t} \cdot p_{1,t+1} \cdots p_{n-1,t+n-1}$. En esta sección, se ha vuelto a ajustar un modelo Lee-Carter binomial y Lee-Carter gaussiano entre los años 1992 y 2008, pero, ahora, se ha extendido el rango de edad hasta los 100 años y se han realizado las proyecciones de probabilidad de supervivencia. En la tabla 9, se presenta la esperanza de vida al nacer, por sexo, para una persona que nace en el año 2019 ($t = 2019$) y para una persona que nace en el año 2050 ($t = 2050$), según los modelos Lee-Carter binomial y Lee-Carter gaussiano.

Tabla 9. Esperanza de vida pronosticada para el año 2019 y el año 2050, según sexo.

Modelos	Año 2019		Año 2050	
	Hombre	Mujer	Hombre	Mujer
Lee-Carter Binomial	85,05	90,26	87,95	92,47
Lee-Carter Binomial-Low	86,41	90,80	89,04	92,87
Lee-Carter Binomial-Up	83,45	89,68	86,59	92,01
Lee-Carter Gaussiano	84,48	89,56	87,99	91,65
Lee-Carter Gaussiano-Low	85,71	90,72	88,16	92,55
Lee-Carter Gaussiano-Up	88,02	88,02	85,93	90,44

Tomando como base las predicciones realizadas por el modelo Lee-Carter Binomial, en el 2050 la mujer al nacer tendrá una esperanza de vida de 92,47 años y el hombre una esperanza de vida de 87,95 años. Las diferencias en la mortalidad por sexo se mantendrán en el futuro. En 2050 los hombres y mujeres vivirán en promedio 2,90 y 2,21 años más que en la actualidad, respectivamente. En el caso del modelo Lee-Carter gaussiano, el comportamiento es muy similar, aunque las predicciones de la esperanza de vida futura son ligeramente inferiores.

5. Conclusiones

Dada su utilidad para la planificación de los sistemas de seguridad social y para las compañías privadas de seguros, entre otras materias, es de gran interés contar con modelos que logren describir de forma correcta los datos observados, y, también, es relevante que los modelos generen pronósticos fiables. En este artículo, se ha presentado la familia de modelos de mortalidad estocástica edad-periodo-cohorta generalizados (GAPC, en sus siglas en inglés) de Villegas et al. (2017), que abarca la gran mayoría de modelos de proyección de mortalidad. Asimismo, se han presentado los modelos de Lee-Carter, Renshaw-Haberman y Cairns-Blake-Dowd, expuestos bajo el marco de la familia de modelos GAPC, así como, también, el modelo de Lee-Carter en su versión original. Se han utilizado estos modelos para ajustar datos de mortalidad para los hombres y mujeres de Chile, considerando el rango de edad que va de 0 a 90 años, y para los años calendario de 1992 a 2008 (datos que han sido obtenidos de la base de datos The Human Mortality Database).

Se han comparado los ajustes realizados con los modelos GAPC utilizando como criterio las medidas de bondad de ajuste AIC y BIC, asumiendo distribución Binomial y distribución Poisson para el número de muertes. En el caso del modelo Lee-Carter y Renshaw-Haberman, para ambos sexos, se obtienen mejores resultados utilizando distribución Binomial, mientras que para el modelo de Cairns-Blake-Dowd se obtiene un mejor ajuste cuando se asume distribución Poisson. Sin embargo, en base al análisis de los residuos, constatamos que el modelo de Cairns-Blake-Dowd no presenta un adecuado ajuste de los datos.

También se han comparado los ajustes utilizando como criterio los valores del error cuadrático medio y del error porcentual absoluto medio, entre el número de muertes observadas y estimadas, para los modelos Lee-Carter y Renshaw-Haberman con distribución binomial y el modelo CBD con Poisson, así como para el modelo Lee Carter original. Tanto para hombres como para mujeres, el modelo de Renshaw-Haberman es el que presenta menor error de estimación, con una leve diferencia con respecto a los desempeños de los modelos Lee-Carter binomial y gaussiano, mientras que el modelo Cairns-Blake-Dowd es el que presenta los peores resultados.

Se generaron proyecciones futuras para las tasas de mortalidad, y con estos datos se ha calculado el error cuadrático medio y el error porcentual absoluto medio entre las tasas de mortalidad observadas por el Instituto Nacional de Estadísticas de Chile para los años 2009 a 2012 y las tasas de mortalidad

pronosticadas por los modelos de Lee-Carter gaussiano y binomial, y por el modelo de Renshaw-Haberman binomial, para los mismos años. Se obtuvo, tanto para hombres como para mujeres, que el modelo de Lee-Carter es el que presentó menor error de predicción (especialmente, el basado en la distribución binomial).

Con los pronósticos realizados, se construyeron tablas de mortalidad para el cálculo de primas en seguros de vida, en dos escenarios hipotéticos. Estas primas, se compararon a su vez con las obtenidas utilizando las tablas de mortalidad M95, que son las tablas de mortalidad publicadas por el ente regulador de Chile para la constitución de reservas de pólizas de seguros de vida. En vista de los resultados, las tablas M95 se presentan conservadoras en comparación a los otros modelos, y parecen adecuadas para garantizar la solvencia de las entidades aseguradoras, en cuanto a posibles desviaciones en el riesgo de fallecimiento.

En definitiva, para los datos utilizados en este estudio, hemos constatado que el modelo de Renshaw-Haberman es el que logra un mejor ajuste a los datos observados, pero es menos útil para pronosticar. Por lo tanto, considerando el desempeño, tanto para describir el pasado como para generar pronósticos, y también por su simplicidad, concluimos que el modelo de Lee-Carter Binomial es el más adecuado para describir el comportamiento y la evolución de la mortalidad de la población de Chile. Con este modelo, se calculó la esperanza de vida al nacer en el año 2019 y el año 2050 para hombres y mujeres en Chile. En el año 2050, la esperanza de vida al nacer estará en promedio por encima de los 90 años en Chile (92,47 años para las mujeres y 87,95 años para los hombres), lo que plantea enormes desafíos a los sistemas de previsión públicos y privados.

La principal conclusión del estudio que debemos destacar, es que la longevidad de la población chilena se espera que siga aumentando en los próximos años. Las tablas de mortalidad que actualmente utilizan las compañías aseguradoras chilenas en el cálculo de las reservas en los seguros de vida parecen adecuadas para garantizar su suficiencia. Estas tablas están vigentes desde el año 2000 y no tienen en cuenta la evolución de la mortalidad de la población chilena. Si las tasas de mortalidad continúan disminuyendo, sería recomendable que estas tablas fueran revisadas en los próximos años para adaptarlas a esta circunstancia.

La utilización de tablas de mortalidad generacionales que recojan la evolución futura de la mortalidad es más importante, si cabe, para el cálculo de las pensiones y rentas temporales y vitalicias. La Superintendencia de Pensiones

debe, por ley, establecer y actualizar periódicamente las tablas de mortalidad que se utilizan para el cálculo de las pensiones por retiro programado. Estas tablas estimaban una esperanza de vida de 85,3 años para un hombre que en el año 2016 tuviera 65 años y una esperanza de vida de 90,4 años para una mujer que en ese mismo año tuviera 60 años (Macias, 2016). Las tablas de mortalidad de este artículo se basan en la población general. Por el contrario, los datos utilizados por la Superintendencia de Pensiones son de la subpoblación de pensionados que financian una pensión mayor a la pensión básica solidaria. Aunque la propia Superintendencia señala que se deben comparar con cuatela las tablas de mortalidad para toda la población con las tablas de mortalidad para pensionados, los resultados mostrados en este artículo permiten hacerse una idea sobre la evolución esperada de la mortalidad en la población chilena de pensionados para los próximos años.

El estudio presenta algunas limitaciones que deben ser mencionadas. Una de las principales limitantes presentadas en el desarrollo del estudio han sido los datos. Si bien es cierto que la información de la mortalidad en Chile es de buena calidad, en la actualidad la información de público acceso es limitada. Los datos disponibles en la base de datos The Human Mortality Database cubren únicamente el periodo temporal de 1992 hasta 2008 y sería deseable disponer de información más reciente para la calibración de los modelos estadísticos. Por otro lado, la información de público acceso en el INE de Chile dispone las edades agrupadas en formato 0, 1-4, 5-9,... 75-79, 80+ y dispone los años de calendario en quinquenios, lo que se traduce en una menor cantidad de observaciones para efectos de la estimación de los parámetros de cada modelo. En este sentido, futuros trabajos académicos podrían abocarse a la recopilación y/o desarrollo mediante técnicas de interpolación de una mayor cantidad de observaciones, lo que a su vez permitiría desarrollar una comparación con más modelos de mortalidad que los que se han considerado en este estudio.

Referencias



- Arnold, M., Herrera, F., Massad, C. y Thumala, D. (2018). *Quinta encuesta nacional de inclusión y exclusión social de las personas mayores de Chile: opiniones de la población chilena respecto al envejecimiento poblacional*. Santiago: Ediciones Servicio Nacional del Adulto Mayor.
- Ayuso, M. . Bravo, J. y Holzmänn. R. (2020). Getting life expectancy estimates right for pension policy: period versus cohort approach. *Journal of Pension Economics and Finance*, (en prensa).

- Ayuso, M., Corrales, H., Guillén, M., Pérez-Marín, A. M. & Rojo, J. L. (2007). *Estadística Actuarial Vida* (7 edición). Barcelona: Ediciones UB.
- Betzuen, A. (2010). Un análisis sobre las posibilidades de predicción de la mortalidad futura aplicando el modelo lee-carter. *Anales del Instituto de Actuarios Españoles*, 16, 111–140.
- Blake, D., Cairns, A. y Dowd, K. (2006). A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *Journal of Risk & Insurance*, 73, 687–718.
- Brouhns, N., Denuit, M. y Vermunt, J. K. (2002). A poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics*, 31(3), 373–393.
- CEPAL (2017). *Tablas de mortalidad (LC/PUB.2017/20-P)*. Santiago: Comisión Económica para América Latina y el Caribe.
- CEPAL (2000). *El envejecimiento de la población latinoamericana: ¿hacia una relación de dependencia favorable?* Santiago: Centro Latinoamericano y Caribeño de Demografía (CELADE) – División de Población.
- Currie, I. D. (2016). On fitting generalized linear and non-linear models of mortality. *Scandinavian Actuarial Journal*, 4, 356–383.
- Debón, A., Martínez, F., Montes, F. Y Moshuk, M. (2015). Proyección de los indicadores de mortalidad para España. *Estadística Española*, 57(188), 291–321.
- DEIS (2020). *Series y gráficos de natalidad*. Santiago: Departamento de Estadísticas e Información de Salud, Ministerio de Salud, Gobierno de Chile.
- Haberman, S. y Renshaw, A. (2011). A comparative study of parametric mortality projection models. *Insurance: Mathematics and Economics*, 48(1), 35–55.
- Jiménez, A. (2008). *Hacia un sistema Nacional de Promoción de la Autonomía Personal y Atención a las Personas en situación de Dependencia. Condicionantes demográficos, modelos internacionales*,

propuestas de lineamientos de políticas y recomendaciones. Santiago: SENADIS, Ministerio de Desarrollo Social y Familia.

- Lee, R.D. y Rofman, R. (1994). Modeling and projecting mortality in Chile. *Notas de población*, 22(59), 183–213.
- Lee, R. D. y Carter, L. R. (1992). Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association*, 87(419), 659–671.
- Macias, O. (2016) *Nuevas tablas de mortalidad del sistema de pensiones: Comisión de familia y adulto mayor*, Superintendencia de pensiones (Consulta: Noviembre, 2020 https://www.spensiones.cl/portal/institucional/594/articles-10993_recurso_1.pdf)
- Macias, Y. y Santolino, M. (2018). Aplicación de los modelos Lee-Carter y Renshaw-Haberman en los seguros de vida y mixtos. *Anales del Instituto de Actuarios Españoles*, 4(24), 53–78.
- Melgen-Bello, L. y García-Prieto, C. (2017). Análisis de la eficiencia del gasto sanitario de los países de América Latina y el Caribe. *Salud Publica de Mexico*, 583–591.
- Plat, R. (2009). On stochastic mortality modeling. *Insurance: Mathematics and Economics*, 45(3), 393–404.
- Renshaw, A.E. y Haberman, S. (2003). Lee-Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics*, 33(2), 255–272.
- Renshaw, A.E. y Haberman, S. (2006). A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, 38(3), 556–570.
- R Core Team. (2020). *R: A Language and Environment for Statistical Computing*, 7 edition. Vienna: R Foundation for Statistical Computing.
- Santolino, M. (2020). The lee-carter quantile mortality model. *Scandinavian Actuarial Journal*, (En prensa).
- Villegas, A. M., Millossovich, P. y Kaishev, V. K. (2017). StMoMo: Stochastic mortality modeling in r. *Journal of Statistical Software*,

84(3), 1–38.